

09/462625
11 JAN 2000

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: Boehringer Ingelheim International GmbH
(B) STREET: Binger Strasse 173
(C) CITY: Ingelheim am Rhein
(E) COUNTRY: Germany
(F) POSTAL CODE (ZIP): 55216
(G) TELEPHONE: 06132/772282
(H) TELEFAX: 06132/774377

(ii) TITLE OF INVENTION: Tumor Growth Inhibition- and
Apoptosis-Associated Genes and Polypeptides and
Methods of Use Thereof

(iii) NUMBER OF SEQUENCES: 24

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(v) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 893 764
(B) FILING DATE: 11-JUL-1997

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 549 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 1..549

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

5	ATG TTG TTT GCC TGT GCT CTC CTT GCC CTC CTG GGT CTG GCA ACC TCC	48
	Met Leu Phe Ala Cys Ala Leu Leu Ala Leu Leu Gly Leu Ala Thr Ser	
	1 5 10 15	
10	TGC AGT TTC ATC GTG CCC CGC AGT GAG TGG AGG GCC CTG CCA TCC GAG	96
	Cys Ser Phe Ile Val Pro Arg Ser Glu Trp Arg Ala Leu Pro Ser Glu	
	20 25 30	
15	TGC TCT AGC CGC CTG GGG CAC CCA GTT CGC TAC GTG GTG ATC TCA CAC	144
	Cys Ser Ser Arg Leu Gly His Pro Val Arg Tyr Val Val Ile Ser His	
	35 40 45	
20	ACA GCC GGC AGC TTC TGC AAC AGC CCG GAC TCC TGT GAA CAG CAG GCC	192
	Thr Ala Gly Ser Phe Cys Asn Ser Pro Asp Ser Cys Glu Gln Gln Ala	
	50 55 60	
25	CGC AAT GTG CAG CAT TAC CAC AAG AAT GAG CTG GGC TGG TGC GAT GTA	240
	Arg Asn Val Gln His Tyr His Lys Asn Glu Leu Gly Trp Cys Asp Val	
	65 70 75 80	
30	GCC TAC AAC TTC CTT ATT GGA GAG GAC GGT CAT GTC TAT GAA GGC CGA	288
	Ala Tyr Asn Phe Leu Ile Gly Glu Asp Gly His Val Tyr Glu Gly Arg	
	85 90 95	
35	GGC TGG AAC ATC AAG GGT GAC CAC ACA GGG CCC ATC TGG AAT CCC ATG	336
	Gly Trp Asn Ile Lys Gly Asp His Thr Gly Pro Ile Trp Asn Pro Met	
	100 105 110	
40	TCT ATT GGC ATC ACC TTC ATG GGG AAC TTC ATG GAC CGG GTA CCC GCA	384
	Ser Ile Gly Ile Thr Phe Met Gly Asn Phe Met Asp Arg Val Pro Ala	
	115 120 125	
45	AAG CGG GCC CTC CGT GCT GCC CTA AAT CTT CTG GAA TGT GGG GTG TCT	432
	Lys Arg Ala Leu Arg Ala Ala Leu Asn Leu Leu Glu Cys Gly Val Ser	
	130 135 140	
50	CGG GGC TTC CTG AGA TCC AAC TAT GAA GTC AAA GGA CAC CGG GAT GTG	480
	Arg Gly Phe Leu Arg Ser Asn Tyr Glu Val Lys Gly His Arg Asp Val	
	145 150 155 160	
55	CAA AGC ACT CTC TCT CCA GGT GAC CAA CTC TAT CAG GTC ATC CAA AGC	528
	Gln Ser Thr Leu Ser Pro Gly Asp Gln Leu Tyr Gln Val Ile Gln Ser	
	165 170 175	
60	TGG GAA CAC TAC CGA GAG TGA	549
	Trp Glu His Tyr Arg Glu	
	180	

009162625-072800

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 182 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Leu Phe Ala Cys Ala Leu Leu Ala Leu Leu Gly Leu Ala Thr Ser
 1 5 10 15

Cys Ser Phe Ile Val Pro Arg Ser Glu Trp Arg Ala Leu Pro Ser Glu
 20 25 30

Cys Ser Ser Arg Leu Gly His Pro Val Arg Tyr Val Val Ile Ser His
 35 40 45

Thr Ala Gly Ser Phe Cys Asn Ser Pro Asp Ser Cys Glu Gln Gln Ala
 50 55 60

Arg Asn Val Gln His Tyr His Lys Asn Glu Leu Gly Trp Cys Asp Val
 65 70 75 80

Ala Tyr Asn Phe Leu Ile Gly Glu Asp Gly His Val Tyr Glu Gly Arg
 85 90 95

Gly Trp Asn Ile Lys Gly Asp His Thr Gly Pro Ile Trp Asn Pro Met
 100 105 110

Ser Ile Gly Ile Thr Phe Met Gly Asn Phe Met Asp Arg Val Pro Ala
 115 120 125

Lys Arg Ala Leu Arg Ala Ala Leu Asn Leu Leu Glu Cys Gly Val Ser
 130 135 140

Arg Gly Phe Leu Arg Ser Asn Tyr Glu Val Lys Gly His Arg Asp Val
 145 150 155 160

Gln Ser Thr Leu Ser Pro Gly Asp Gln Leu Tyr Gln Val Ile Gln Ser
 165 170 175

Trp Glu His Tyr Arg Glu
 180

008220-52929460

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 718 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens
 (F) TISSUE TYPE: Bone Marrow

(ix) FEATURE:

- (A) NAME/KEY: 5'UTR
 (B) LOCATION:1..67

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION:68..643

(ix) FEATURE:

- (A) NAME/KEY: 3'UTR
 (B) LOCATION:644..718

(ix) FEATURE:

- (A) NAME/KEY: polyA_site
 (B) LOCATION:712..714

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

CTGAGTTACT GGGCCCAGAG GCTGGGCCCC TGGACATGTA CCTGCAGCCA CTATGTCCCG      60
40 CCGCTCT ATG CTG CTT GCC TGG GCT CTC CCC AGC CTC CTT CGA CTC GGA      109
    Met Leu Leu Ala Trp Ala Leu Pro Ser Leu Leu Arg Leu Gly
      1              5              10

GCG GCT CAG GAG ACA GAA GAC CCG GCC TGC TGC AGC CCC ATA GTG CCC      157
45 Ala Ala Gln Glu Thr Glu Asp Pro Ala Cys Cys Ser Pro Ile Val Pro
   15              20              25              30

CGG AAC GAG TGG AAG GCC CTG GCA TCA GAG TGC GCC CAG CAC CTG AGC      205
50 Arg Asn Glu Trp Lys Ala Leu Ala Ser Glu Cys Ala Gln His Leu Ser
      35              40              45

CTG CCC TTA CGC TAT GTG GTG GTA TCG CAC ACG GCG GGC AGC AGC TGC      253
    Leu Pro Leu Arg Tyr Val Val Val Ser His Thr Ala Gly Ser Ser Cys
      50              55              60

AAC ACC CCC GCC TCG TGC CAG CAG CAG GCC CGG AAT GTG CAG CAC TAC      301
55 Asn Thr Pro Ala Ser Cys Gln Gln Gln Ala Arg Asn Val Gln His Tyr
      65              70              75

```

008220" 52929460

5	CAC ATG AAG ACA CTG GGC TGG TGC GAC GTG GGC TAC AAC TTC CTG ATT His Met Lys Thr Leu Gly Trp Cys Asp Val Gly Tyr Asn Phe Leu Ile 80 85 90	349
10	GGA GAA GAC GGG CTC GTA TAC GAG GGC CGT GGC TGG AAC TTC ACG GGT Gly Glu Asp Gly Leu Val Tyr Glu Gly Arg Gly Trp Asn Phe Thr Gly 95 100 105 110	397
15	GCC CAC TCA GGT CAC TTA TGG AAC CCC ATG TCC ATT GGC ATC AGC TTC Ala His Ser Gly His Leu Trp Asn Pro Met Ser Ile Gly Ile Ser Phe 115 120 125	445
20	ATG GGC AAC TAC ATG GAT CGG GTG CCC ACA CCC CAG GCC ATC CGG GCA Met Gly Asn Tyr Met Asp Arg Val Pro Thr Pro Gln Ala Ile Arg Ala 130 135 140	493
25	GCC CAG GGT CTA CTG GCC TGC GGT GTG GCT CAG GGA GCC CTG AGG TCC Ala Gln Gly Leu Leu Ala Cys Gly Val Ala Gln Gly Ala Leu Arg Ser 145 150 155	541
30	AAC TAT GTG CTC AAA GGA CAC CGG GAT GTG CAG CGT ACA CTC TCT CCA Asn Tyr Val Leu Lys Gly His Arg Asp Val Gln Arg Thr Leu Ser Pro 160 165 170	589
35	GGC AAC CAG CTC TAC CAC CTC ATC CAG AAT TGG CCA CAC TAC CGC TCC Gly Asn Gln Leu Tyr His Leu Ile Gln Asn Trp Pro His Tyr Arg Ser 175 180 185 190	637
40	CCC TGA GGCCCTGCTG ATCCGCACCC CATTCCTCCC CTCCCATGGC CAAAAACCCC Pro *	693
45	ACTGTCTCCT TCTCCAATAA AGATG	718

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 192 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met Leu Leu Ala Trp Ala Leu Pro Ser Leu Leu Arg Leu Gly Ala Ala
1 5 10 15

Gln Glu Thr Glu Asp Pro Ala Cys Cys Ser Pro Ile Val Pro Arg Asn
20 25 30

Glu Trp Lys Ala Leu Ala Ser Glu Cys Ala Gln His Leu Ser Leu Pro
35 40 45

09462625 "072800

Leu Arg Tyr Val Val Val Ser His Thr Ala Gly Ser Ser Cys Asn Thr
 50 55 60
 5 Pro Ala Ser Cys Gln Gln Gln Ala Arg Asn Val Gln His Tyr His Met
 65 70 75 80
 Lys Thr Leu Gly Trp Cys Asp Val Gly Tyr Asn Phe Leu Ile Gly Glu
 85 90 95
 10 Asp Gly Leu Val Tyr Glu Gly Arg Gly Trp Asn Phe Thr Gly Ala His
 100 105 110
 Ser Gly His Leu Trp Asn Pro Met Ser Ile Gly Ile Ser Phe Met Gly
 115 120 125
 Asn Tyr Met Asp Arg Val Pro Thr Pro Gln Ala Ile Arg Ala Ala Gln
 130 135 140
 20 Gly Leu Leu Ala Cys Gly Val Ala Gln Gly Ala Leu Arg Ser Asn Tyr
 145 150 155 160
 Val Leu Lys Gly His Arg Asp Val Gln Arg Thr Leu Ser Pro Gly Asn
 165 170 175
 25 Gln Leu Tyr His Leu Ile Gln Asn Trp Pro His Tyr Arg Ser Pro *
 180 185 190

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TTTTTTTTTT TTAC

14

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

008270" 52929460

(ii) MOLECULE TYPE: synthetic DNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AATCGGGCTG

10

10

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

15

(A) LENGTH: 10 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

25

AGTCAGCCAC

10

30

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

35

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GAAAGGGCCT CGTGATAC

18

45

(2) INFORMATION FOR SEQ ID NO:9:

50

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

09462625.072800

(ii) MOLECULE TYPE: synthetic DNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CAGGGCCCTC CACTCACTTG CATAGGCATT TGTAGC

36

10

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

25 GCTACAAATG CCTATGCAAG TGAGTGGAGG GCCCTG

36

25

(2) INFORMATION FOR SEQ ID NO:11:

30

(i) SEQUENCE CHARACTERISTICS:

- 35 (A) LENGTH: 41 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGCCGCTAGC CTGCAGTTAT CACTCTCGGT AGTGTTCCCA G

41

45

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

5

GGCGGATCCG AGTGGAGGGC CCTGCCATCC

30

10

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

15

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

25

GGCGAATTCT TATCACTCTC GGTAGTGTTT

30

30

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

35

(A) LENGTH: 36 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

40

(ii) MOLECULE TYPE: synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GGCCAGATCT CGTCCAGCAT GTTGTGTTGCC TGTGCT

36

45

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

50

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

008220"52929460

(ii) MOLECULE TYPE: synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GGCACCAAAA TCAACGGGAC

20

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CAGGGCCCTC CACTCACATC GTGCACCTGG G

31

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CCCAGGTGCA CGATGTGAGT GGAGGGCCCT G

31

(2) INFORMATION FOR SEQ ID NO:18

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GGCCGCTAGC CTGCAGTTAT CACTCTCGGT AGTGTTCCTCA G

41

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GGCCAAGCTT CCACCATGTC CCGCCGCTCT ATGCTG

36

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GGCCGAATTC TTATCAGGGG GAGCGGTAGT GTGG

34

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

5 Arg Ser Glu Trp Arg Ala Leu Pro Ser Glu Cys Ser Ser Arg Leu Gly
 1 5 10 15
 His Pro Tyr

10

(2) INFORMATION FOR SEQ ID NO:22:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

25 Gly Glu Asp Gly His Val Tyr Glu Gly Arg Gly Trp Asn Ile Lys Gly
 1 5 10 15
 Asp His Thr Gly Cys Tyr
 30 20

(2) INFORMATION FOR SEQ ID NO:23:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 26 amino acids
 (B) TYPE: amino acid
 40 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

45 Ala Gln Glu Thr Glu Asp Pro Ala Cys Cys Ser Pro Ile Val Pro Arg
 1 5 10 15
 Asn Glu Trp Lys Ala Leu Ala Ser Glu Tyr
 50 20 25

008220" 52929460

J)

(2) INFORMATION FOR SEQ ID NO:24:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

15

Gly Glu Asp Gly Leu Val Tyr Glu Gly Arg Gly Trp Asn Phe Thr Gly
1 5 10 15

Cys Tyr

009462625-072800

-1-

SEQUENCE LISTING

<110> Georgiev, Georgii
Kiselev, Sergei
Prokhorchouk, Egor
Ostermann, Elinborg

<120> Tumor Growth Inhibition- and Apoptosis-Associated Genes
and Polypeptides and Methods of Use Thereof

<130> 0652.1630001

<140>

<141>

<150> US 08/893,764

<151> 1997-07-11

<150> PCT/EP98/04287

<151> 1998-07-10

<160> 24

<170> PatentIn Ver. 2.1

<210> 1

<211> 549

<212> DNA

<213> Artificial Sequence

<220>

<221> CDS

<222> (1)..(549)

<220>

<223> Description of Artificial Sequence: tag7 cDNA

<400> 1
atg ttg ttt gcc tgt gct ctc ctt gcc ctc ctg ggt ctg gca acc tcc 48
Met Leu Phe Ala Cys Ala Leu Leu Ala Leu Leu Gly Leu Ala Thr Ser
1 5 10 15
tgc agt ttc atc gtg ccc cgc agt gag tgg agg gcc ctg cca tcc gag 96
Cys Ser Phe Ile Val Pro Arg Ser Glu Trp Arg Ala Leu Pro Ser Glu
20 25 30
tgc tct agc cgc ctg ggg cac cca gtt cgc tac gtg gtg atc tca cac 144
Cys Ser Ser Arg Leu Gly His Pro Val Arg Tyr Val Val Ile Ser His
35 40 45
aca gcc ggc agc ttc tgc aac agc ccg gac tcc tgt gaa cag cag gcc 192
Thr Ala Gly Ser Phe Cys Asn Ser Pro Asp Ser Cys Glu Gln Gln Ala
50 55 60
cgc aat gtg cag cat tac cac aag aat gag ctg ggc tgg tgc gat gta 240
Arg Asn Val Gln His Tyr His Lys Asn Glu Leu Gly Trp Cys Asp Val
65 70 75 80
gcc tac aac ttc ctt att gga gag gac ggt cat gtc tat gaa ggc cga 288
Ala Tyr Asn Phe Leu Ile Gly Glu Asp Gly His Val Tyr Glu Gly Arg
85 90 95

00462625-072800

ggc tgg aac atc aag ggt gac cac aca ggg ccc atc tgg aat ccc atg 336
Gly Trp Asn Ile Lys Gly Asp His Thr Gly Pro Ile Trp Asn Pro Met 110
100 105

tct att ggc atc acc ttc atg ggg aac ttc atg gac cgg gta ccc gca 384
Ser Ile Gly Ile Thr Phe Met Gly Asn Phe Met Asp Arg Val Pro Ala 125
115 120

aag cgg gcc ctc cgt gct gcc cta aat ctt ctg gaa tgt ggg gtg tct 432
Lys Arg Ala Leu Arg Ala Ala Leu Asn Leu Leu Glu Cys Gly Val Ser 140
130 135

cgg ggc ttc ctg aga tcc aac tat gaa gtc aaa gga cac cgg gat gtg 480
Arg Gly Phe Leu Arg Ser Asn Tyr Glu Val Lys Gly His Arg Asp Val 160
145 150 155

caa agc act ctc tct cca ggt gac caa ctc tat cag gtc atc caa agc 528
Gln Ser Thr Leu Ser Pro Gly Asp Gln Leu Tyr Gln Val Ile Gln Ser 175
165 170

tgg gaa cac tac cga gag tga 549
Trp Glu His Tyr Arg Glu 180

<210> 2

<211> 182

<212> PRT

<213> Artificial Sequence

<223> Description of Artificial Sequence: tag7 cDNA

<400> 2

Met Leu Phe Ala Cys Ala Leu Leu Ala Leu Leu Gly Leu Ala Thr Ser
1 5 10 15
Cys Ser Phe Ile Val Pro Arg Ser Glu Trp Arg Ala Leu Pro Ser Glu
20 25 30
Cys Ser Ser Arg Leu Gly His Pro Val Arg Tyr Val Val Ile Ser His
35 40 45
Thr Ala Gly Ser Phe Cys Asn Ser Pro Asp Ser Cys Glu Gln Gln Ala
50 55 60
Arg Asn Val Gln His Tyr His Lys Asn Glu Leu Gly Trp Cys Asp Val
65 70 75 80
Ala Tyr Asn Phe Leu Ile Gly Glu Asp Gly His Val Tyr Glu Gly Arg
85 90 95
Gly Trp Asn Ile Lys Gly Asp His Thr Gly Pro Ile Trp Asn Pro Met
100 105 110
Ser Ile Gly Ile Thr Phe Met Gly Asn Phe Met Asp Arg Val Pro Ala
115 120 125
Lys Arg Ala Leu Arg Ala Ala Leu Asn Leu Leu Glu Cys Gly Val Ser
130 135 140
Arg Gly Phe Leu Arg Ser Asn Tyr Glu Val Lys Gly His Arg Asp Val
145 150 155 160
Gln Ser Thr Leu Ser Pro Gly Asp Gln Leu Tyr Gln Val Ile Gln Ser
165 170 175
Trp Glu His Tyr Arg Glu 180

<210> 3

<211> 718

<212> DNA

003220"52929460

<213> Homo sapiens

<220>

<221> 5'UTR

<222> (1)..(67)

<220>

<221> CDS

<222> (68)..(643)

<220>

<221> 3'UTR

<222> (644)..(718)

<220>

<221> polyA_site

<222> (712)..(714)

<400> 3

ctgagttact gggcccagag gctgggcccc tggacatgta cctgcagcca ctatgtcccg 60

ccgctct atg ctg ctt gcc tgg gct ctc ccc agc ctc ctt cga ctc gga 109
Met Leu Leu Ala Trp Ala Leu Pro Ser Leu Leu Arg Leu Gly
1 5 10

gcg gct cag gag aca gaa gac ccg gcc tgc tgc agc ccc ata gtg ccc 157
Ala Ala Gln Glu Thr Glu Asp Pro Ala Cys Cys Ser Pro Ile Val Pro
15 20 25 30

cgg aac gag tgg aag gcc ctg gca tca gag tgc gcc cag cac ctg agc 205
Arg Asn Glu Trp Lys Ala Leu Ala Ser Glu Cys Ala Gln His Leu Ser
35 40 45

ctg ccc tta cgc tat gtg gtg gta tgc cac acg gcg ggc agc agc tgc 253
Leu Pro Leu Arg Tyr Val Val Val Ser His Thr Ala Gly Ser Ser Cys
50 55 60

aac acc ccc gcc tcg tgc cag cag cag gcc cgg aat gtg cag cac tac 301
Asn Thr Pro Ala Ser Cys Gln Gln Gln Ala Arg Asn Val Gln His Tyr
65 70 75

cac atg aag aca ctg ggc tgg tgc gac gtg ggc tac aac ttc ctg att 349
His Met Lys Thr Leu Gly Trp Cys Asp Val Gly Tyr Asn Phe Leu Ile
80 85 90

gga gaa gac ggg ctc gta tac gag ggc cgt ggc tgg aac ttc acg ggt 397
Gly Glu Asp Gly Leu Val Tyr Glu Gly Arg Gly Trp Asn Phe Thr Gly
95 100 105 110

gcc cac tca ggt cac tta tgg aac ccc atg tcc att ggc atc agc ttc 445
Ala His Ser Gly His Leu Trp Asn Pro Met Ser Ile Gly Ile Ser Phe
115 120 125

atg ggc aac tac atg gat cgg gtg ccc aca ccc cag gcc atc cgg gca 493
Met Gly Asn Tyr Met Asp Arg Val Pro Thr Pro Gln Ala Ile Arg Ala
130 135 140

gcc cag ggt cta ctg gcc tgc ggt gtg gct cag gga gcc ctg agg tcc 541
Ala Gln Gly Leu Leu Ala Cys Gly Val Ala Gln Gly Ala Leu Arg Ser
145 150 155

aac tat gtg ctc aaa gga cac cgg gat gtg cag cgt aca ctc tct cca 589

003220"5222460


```
<210> 6
<211> 10
<212> DNA
<213> Artificial Sequence
```

<220>
<223> Description of Artificial Sequence: DNA Primer

<400> 6
aatcgggctg 10

<210> 7
<211> 10
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA Primer

<400> 7
agtcagccac 10

<210> 8
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA Primer

<400> 8
gaaagggcct cgtgatac 18

<210> 9
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA Primer

<400> 9
caggccctc cactcacttg cataggcatt tgtagc 36

<210> 10
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA Primer

<400> 10
gctacaaatg cctatgcaag tgagtggagg gccctg 36

<210> 11
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA Primer

09462625-072800

<400> 11
ggccgctagc ctgcagttat cactctcggt agtggtccca g 41

<210> 12
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA Primer

<400> 12
ggcggatccg agtggagggc cctgcatcc 30

<210> 13
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA Primer

<400> 13
ggcgaattct tatcactctc ggtagtgttc 30

<210> 14
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA Primer

<400> 14
ggccagatct cgtccagcat gttgtttgcc tgtgct 36

<210> 15
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA Primer

<400> 15
ggcaccaaaa tcaacgggac 20

<210> 16
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA Primer

<400> 16
cagggcctc cactcacatc gtgcacctgg g 31

008220"52929460

<210> 17
 <211> 31
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: DNA Primer

<400> 17
 cccaggtgca cgatgtgagt ggagggccct g 31

<210> 18
 <211> 41
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: DNA Primer

<400> 18
 ggccgctagc ctgcagttat cactctcggt agtgttccca g 41

<210> 19
 <211> 36
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: DNA Primer

<400> 19
 ggccaagctt ccaccatgtc ccgccgctct atgctg 36

<210> 20
 <211> 34
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: DNA Primer

<400> 20
 ggccgaattc ttatcagggg gagcggtagt gtgg 34

<210> 21
 <211> 19
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: mutag7 peptide

<400> 21
 Arg Ser Glu Trp Arg Ala Leu Pro Ser Glu Cys Ser Ser Arg Leu Gly
 1 5 10 15

His Pro Tyr

008220"52929460

<220>
<223> Description of Artificial Sequence: mutag7 peptide

```
<210> 23
<211> 26
<212> PRT
<213> Artificial Sequence
```

<220>
<223> Description of Artificial Sequence: hutag7 peptide

```
<210> 24
<211> 18
<212> PRT
<213> Artificial Sequence
```

<220>
<223> Description of Artificial Sequence: hutag7 peptide

```
<400> 24
Gly Glu Asp Gly Leu Val Tyr Glu Gly Arg Gly Trp Asn Phe Thr Gly
  1             5             10             15
Cys Tyr
```